MDM – CUSTOMER DATA INTEGRATION

Discovery System Version 3 Exercise

Product: MDM5.5 SP05

ALON WEINSTEIN, SOLUTION OFFICE, SAP ISRAEL

KRISTIN PATTERSON, SOLUTION OFFICE, SAP AMERICA

Updated: September 14, 2007 by Kristin Patterson

Based on MDM – Customer Data Integration Exercise for MDM SP04.

Table of Contents

MDM Scenario -- Customer Data Integration

1   Introduction

This scenario walks through the most of the cycle of a customer data integration process. It shows how SAP NetWeaver Master Data Management can be used to consolidate master data from several systems, defining matching strategies and identifying duplicate entries. It does not show the setup of data governance using the SAP NetWeaver Master Data Management Workflow component as it requires Microsoft Visio Professional which is not part of the Discovery System setup.

2   Customer Data Integration Vision

The "Company" wants one location to govern its customer data. Currently, the subsidiaries; CompanyA and CompanyB, each generate and control their own customer data. It was discovered that CompanyA and CompanyB had both identical and unique customers. If the customer data was merged together there would be duplicates.

The "Company" decided to implement the standard SAP NetWeaver MDM Customer repository where all customer data would reside. SAP MDM offered the opportunity to import data from both subsidiaries and also identify duplicate records.

3   Customer Data Integration Steps

4) Repository Setup
- Use existing Customer repository
- Add Status field
- Create remote systems

5) and 6) Add Data
- Manually create – Status codes
- Import Lookup table – Countries
- Import – CompanyA and CompanyB records

7) Setup Matching
- Create transformations
- Create rules
- Create strategies
- Find duplicates

4 Repository Setup

First logon to the Discovery System server, either directly, or through Remote Desktop Connection. As the basis for this scenario, we will use the standard Customer repository that SAP provides as business content for MDM.

Download the required files from SDN under SAP Discovery System > Introductory Exercises > Customer Data Integration with MDM > Download MDM files for this exercise. Extract the ZIP to a temp folder on the server. Copy the file 55310_CUSTOMER.A2A to the archives directory under the MDM server's directory ("D:\Program_Files\SAP MDM5.5\Server\Archives\"). This will make the a2a file (archived repository) available for un-archiving from the MDM Console. Other files in the extract will be referenced later in this exercise. You can access them from your temp folder.

4.1 Unarchiving the Repository

Start the MDM Console and mount the MDM server. You can use "localhost" as the host name.
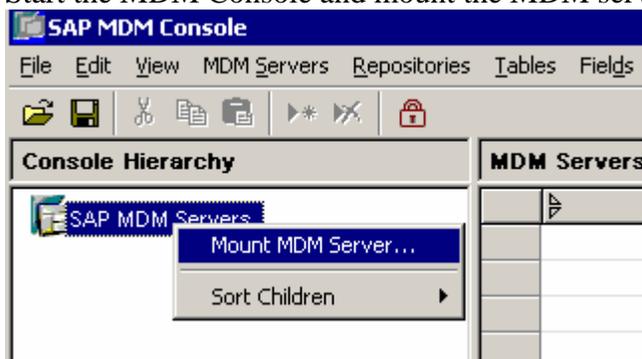
Make sure the MDM server is started – if it is started you will see the following icon next to the hose name: If a different icon is shown right-click the host-name, select "Start MDM Server" and wait for the server to load.

Right-click the host name and select "Unarchive Repository" from the drop-down menu:
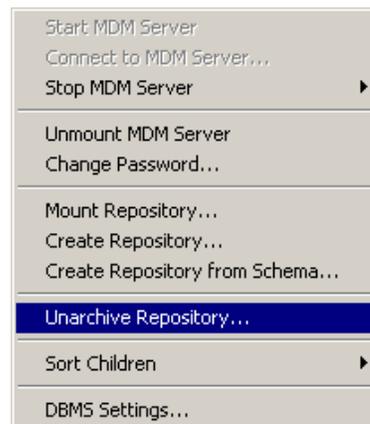
In the pop-up which follows enter select "localhost" as the DBMS server (if needed click the browse "…" button to select it), "DBA" for the username and "admin123" for the password and click "Next >"

Select "55310_CUSTOMER.a2a" from the "Archive" drop-down list and type "ESADS_CDI" as the Repository name. Click "Finish".

MDM will now un-archive the repository, which means the repository's structure and whatever data it holds will be read from the archive file (the .a2a file) and loaded into the database. While the un-archiving process is taking place you will see the following icon in the MDM Console:
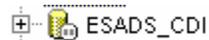

It may take a few minutes until the repository is ready for use. The MDM Console will let you know when the un-archiving is complete with a pop-up message. When you see the following pop-up message click "No" to return to the MDM Console.

You will also notice the icon changes to a grey square icon for repositories that need to be updated. Next, right click and select Update Repository.
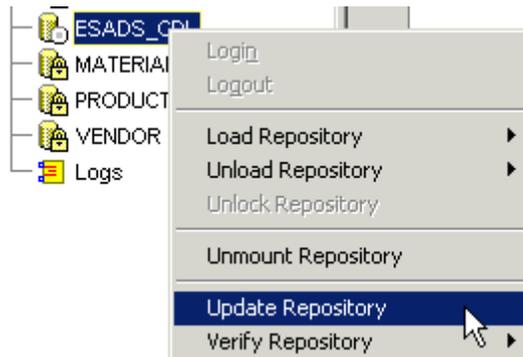

Once the ESADS_CDI repository is updated it will have a lock icon like the rest of the repositories.

Now double-click on the ESADS_CDI rep and login with Username: Admin and password field blank.

4.2 Adding the "Status" Field

To implement the data management process we must add an additional field to the main table. This field will hold information about the current status of a record – is it a newly created record, an imported record, or a confirmed record, which means it is acceptable to syndicate it. Because we need to hold a fixed set of values for this field we will use a flat lookup table to store these values and make the field a flat lookup field in the main table.

Expand the ESADS_CDI repository (keep it unloaded) and click on the Customers main table:
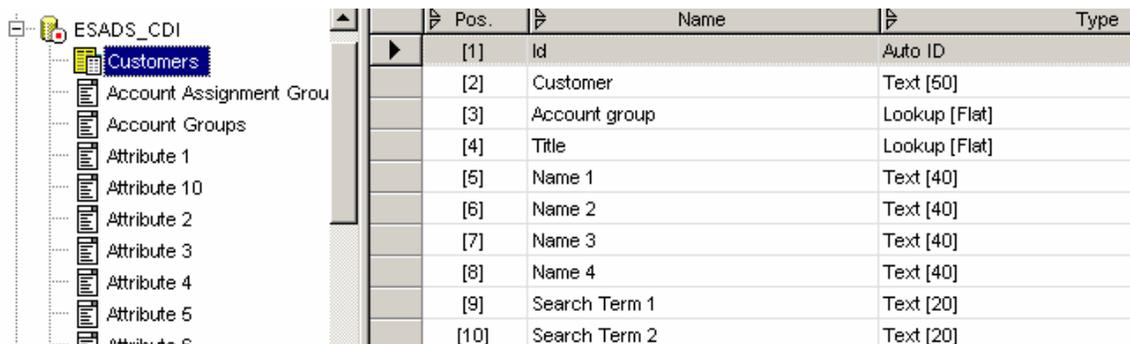
Right-click in the right-hand part of the screen where the table's fields are listed, and select "Add field":

Scroll down to the bottom-right hand part of the screen and enter "Status" as the English name for the field (we will not maintain names in all languages). Click enter to save the name and use the mouse to click on the word "Code" on the left side of the properties panel:

By clicking the "Type" header you will close the multiple-languages input box for the field name and expose the rest of the Field Detail properties input area.

Enter the following values for the field's properties (you might need to scroll down to see all the fields; most of these values are the default values):

Property name	Value
Code	Status
Type	Lookup [Flat]
Required	No
Writable Once	No
Matrix	Normal
Keyword	None
Display field	No
Unique field	No
Sort Index Exists	Normal
Search Tab	Yes
Multi-Valued	No
Lookup Table	[New Flat Table]
New Lookup Table Name	Status Codes

Click Shift+Enter to save the new Status field. MDM will add the new field and also create a new flat lookup table named "Status Codes."

Click on the repository in the left-hand side to see the list of tables and find the new table "Status Codes":

Change the code to "Status_Codes".

Click Shift+Enter to save the changes:

4.3 Creating Remote Systems

Since we will be importing and syndicating data from and to two different systems we want to save remote-key mappings to make sure we do not cause duplicates in the remote systems. To be able to save remote-key mappings we must define Remote systems.

In the left-hand side tree expand the "Admin" node and click "Remote Systems" to access the remote systems administration screen:
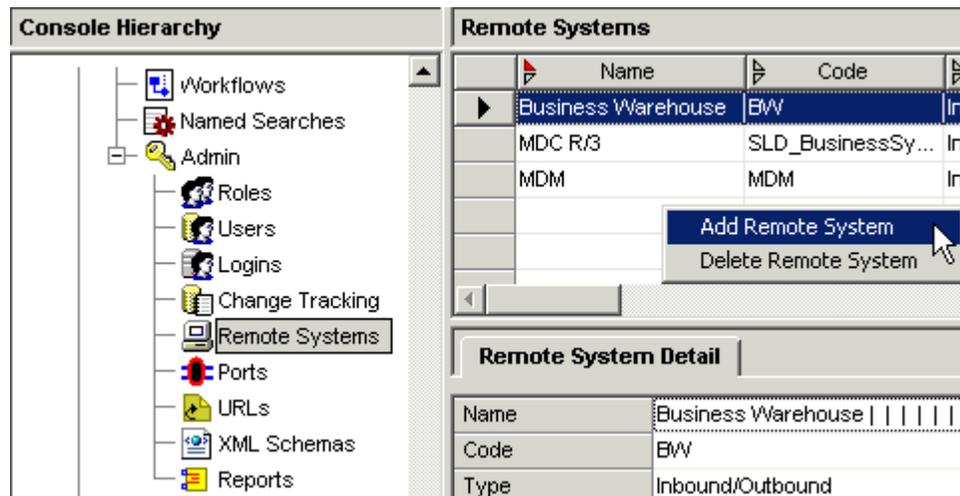
Right-click in the top right-hand side and select "Add Remote System" from the popup down menu.

In the bottom right-hand side of the screen you can input values for the properties of the new remote system. Enter the following values:

Property name	Value
Name	COMPANYA
Code	COMPANYA
Type	Inbound/Outbound
Key Generation	None
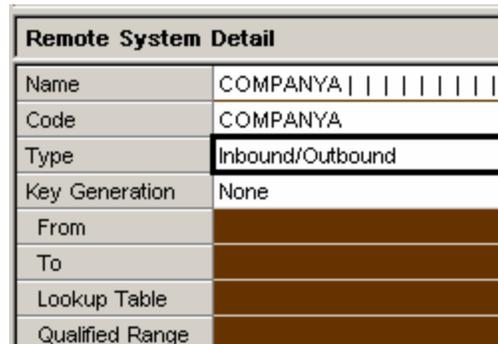
Click Shift+Enter to save.

Repeat this process once more with the following values:

Property name	Value
Name	COMPANYB
Code	COMPANYB
Type	Inbound/Outbound
Key Generation	None

Seen above is the final results after adding the 2 remote systems.

4.4 Loading the Repository

Now that we have the repository configured we can load it for the first time. Right-click the repository and select "Load Repository ▶ Immediate" from the popup menu.


While the repository is loading you will see the following icon: It may take a few minutes to load the repository, especially for the first time after fields and tables were added to allow for internal indexes to be created. Once the repository is loaded the icon will change to: .

5 Adding Status Codes

Open Data Manager and connect to the repository. Select the "ESADS_CDI" repository from the server "localhost", "English [US]" as the language, "Admin" as the username and leave the password field blank:
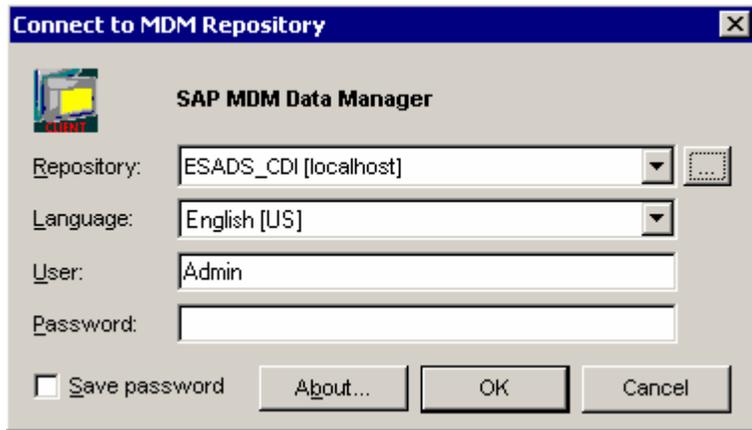

Change the current table to the "Status Codes" table by clicking on the top-left selection box and selecting "Status Codes" from the drop-down list:
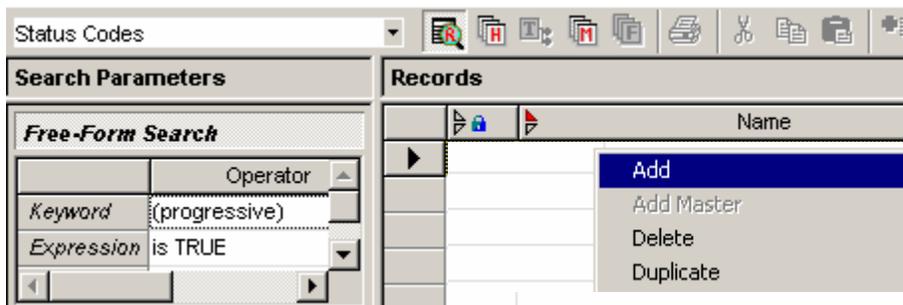
Right-click in the records section pane and select "Add" to create a new record

Type "Approved" as value for the "Name" field and click "Shift+Enter" to save the record. Repeat the process to add two records with the values "Imported" AND "Manually Entered". The result should look like this:
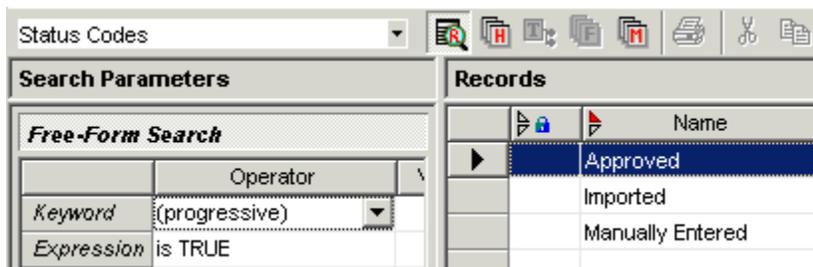
6 Importing Data

We will be importing both lookup table and client data. For the client data, an import map is included in the bundle. MDM comes bundled with import maps to help you quickly import reference data from R/3 systems. We will be using these maps for the lookup tables.

6.1 Updating Remote-key Mappings in the Countries Lookup Table

Before we load data to the main table it would be wise to populate the lookup tables which hold reference data, for example the Countries lookup table. For this scenario we will only preload the Countries lookup table values.

6.1.1 Export Map from Predefined Remote System

The repositories SAP delivers with MDM already have the list of countries populated (including ISO and country codes), but we want to make sure that each country also has a remote-key mapping for the different remote-systems that we are going to use for imports and syndication. This will ensure both auto-mapping of values when importing and correct syndication of country codes. To make things easier the standard repository has an import map available which we can use to quickly update the remote-keys for the various client-systems we plan to use. To use this import map in the new client-systems we will export it from its original MDC R/3 remote-system and import it to the new remote systems we created.

Open Import Manager, select the "ESADS_CDI" repository from the "localhost" server, "English [US]" as the language, "Admin" as the user and leave the password field blank:
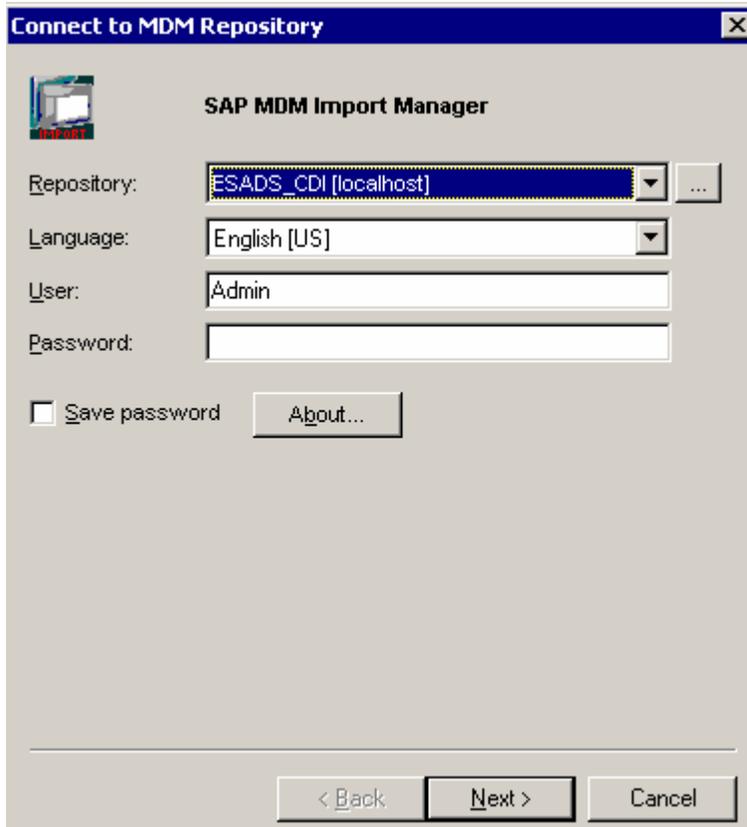
Click "Next >" to continue.

In the next screen select "XML" for Type, "MDC R/3" for Remote system and browse for the file "MDM Lookup Tables.xml" which came with the scenario's bundle.


Click "Finish" to load the file and move on to the Import Manager's main screen.

Click File > Export… and select the map "01_Countries_UpdateRemoteKey_MDM":

Click OK and save the map in the directory where you saved the files from this scenario's bundle. Call the file "UpdateCountriesRemoteKey.map".

## 6.1.2 Import Map into COMPANYA Remote System

Click File > Connect to Source… and change the Remote system to "COMPANYA" and click Finish (do not change any of the other parameters like Type and File name. So use the same values as before):
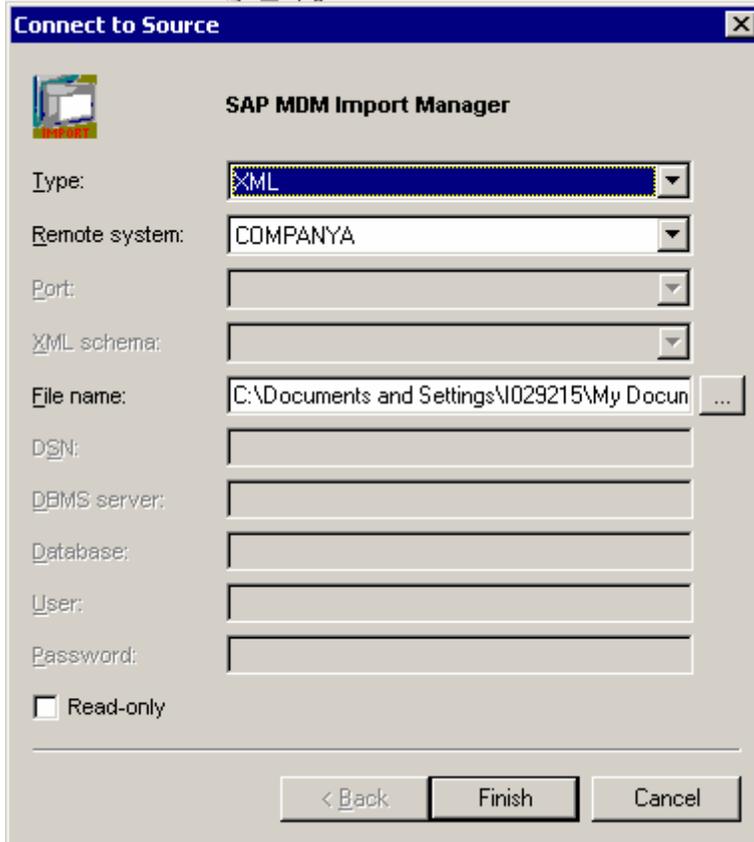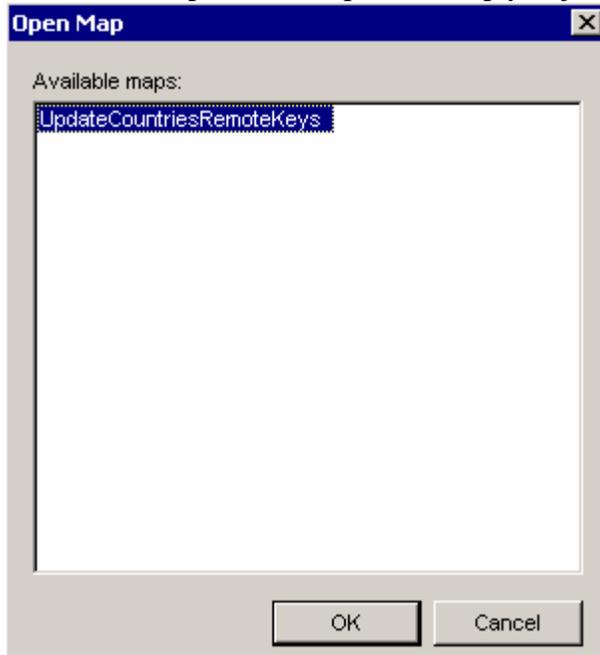

Once the Import Manager is open, click on File > Import… and open the map you saved in the previous step ("UpdateCountriesRemoteKey.map"):
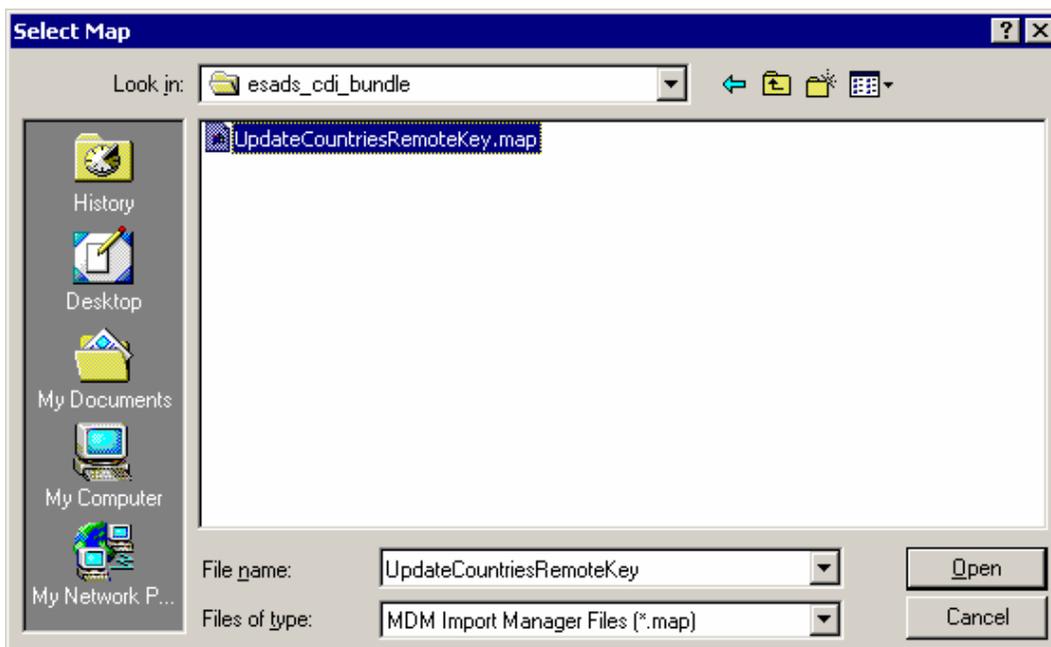
In the "Save Map As" dialog give the map a name, "UpdateCountriesRemoteKeys" and click OK:
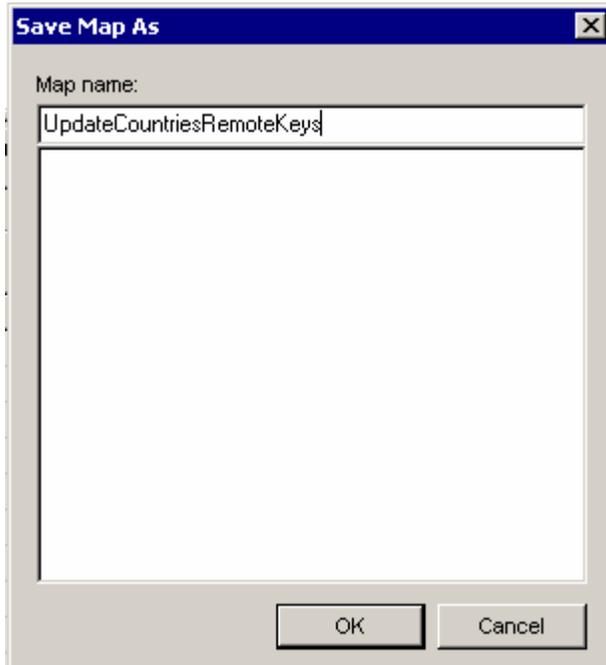
Click File > Open… and open the map you just imported:

You can view the mapping by clicking the "Map Fields/Values" tab. What the map does is add or update to only the remote-key mappings for the current remote-system records in the Countries table.

Click on the "Import Status" and then on the import button in the toolbar to start the import process. MDM will notify you of the import progress and let you know when the process is complete:


6.1.3 Import Map into COMPANYB Remote System

Repeat the same process as the previous step only select "COMPANYB" in the "Connect to Source" dialog box:
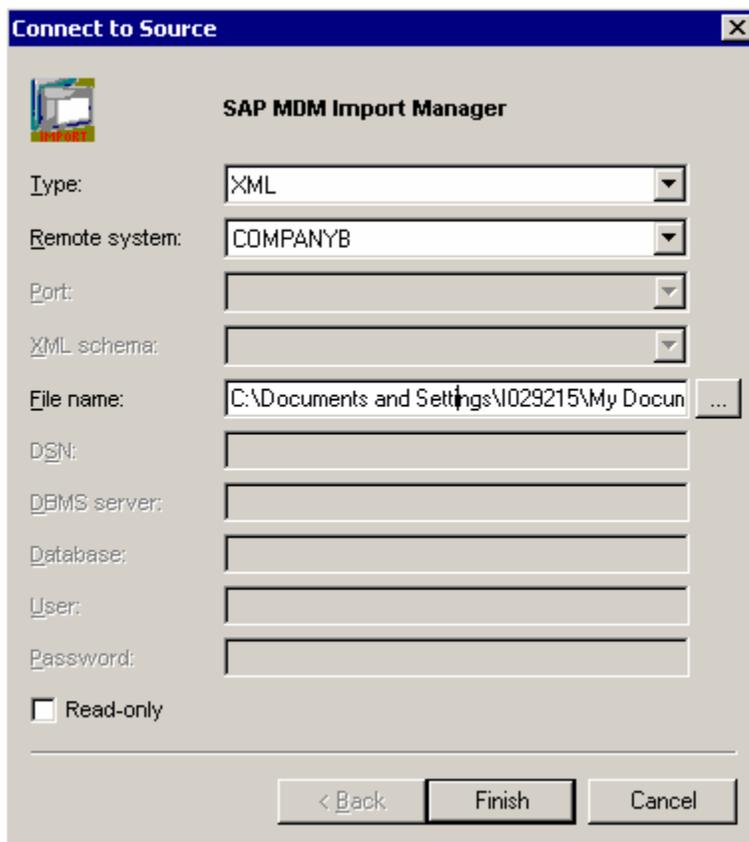
6.1.4 Validating the Import was Successful

To view the imported data, you must open the Data Manager. Login to the repository using the same connection details you used for Import Manager. Change the current table to the "Countries" table by clicking on the top-left selection box and selecting "Countries" from the drop-down list:
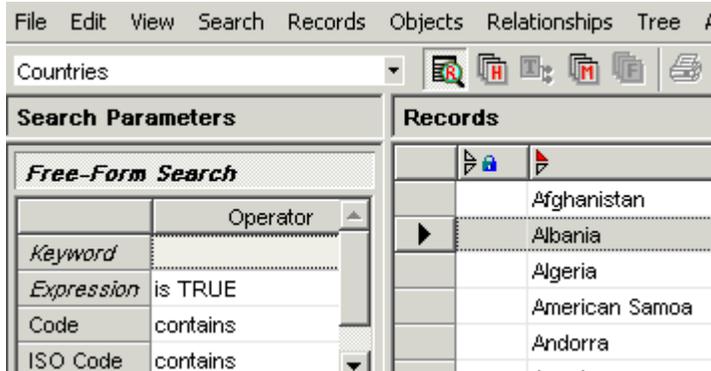
Right-click on any record and select "Edit Key Mappings". You will see that the record has key-mapping information for both Remote Systems; COMPANYA and COMPANYB:
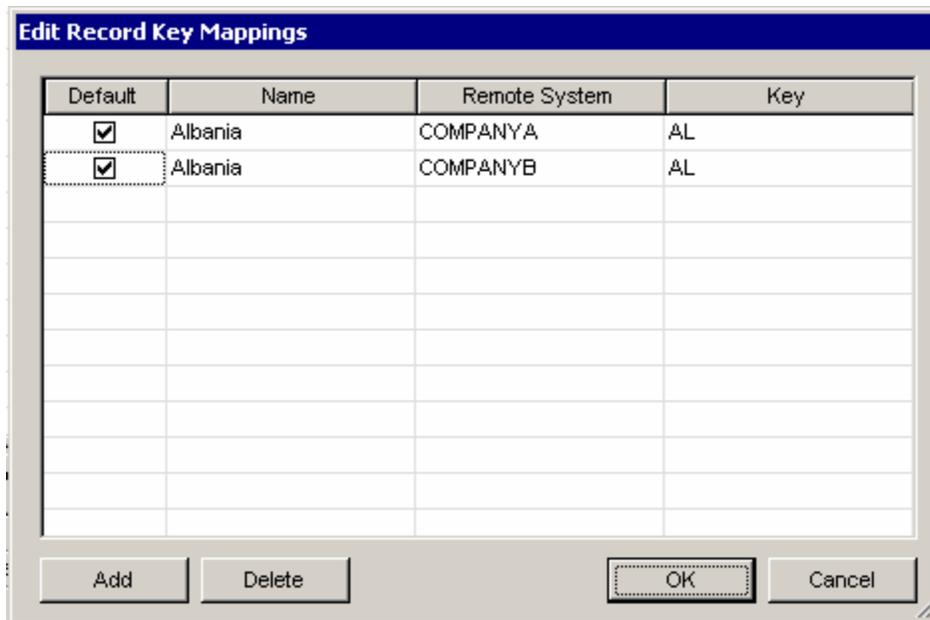
6.2 Loading Data into the Main Table

6.2.1 Loading Data from the COMPANYA Remote System

Open Import Manager and log-in to the repository. In the "Connect to Source" dialog select "Excel" as Type, "COMPANYA" as Remote system and open the file "companya_customers.xls" which came with the scenario's bundle.

Click Finish to continue to Import Manager main screen. Import the map "CustomersTable_SP05.map" which comes with the bundle into the COMPANYA remote system – click File > Import…, open the file and save it in the repository under the name "CustomersTable". Click File > Open… and open the import map you just imported. You can view the mapping by clicking the Map Fields/Values tab:
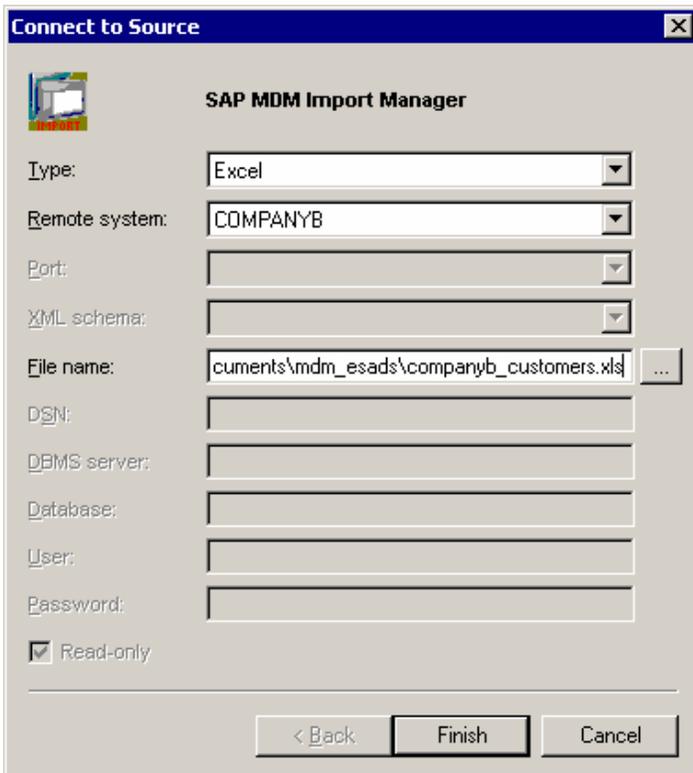
The import map will map most of the fields ("Birthday" and "Gender" are not mapped, as they are not required by the repository's design). In addition, the import map also:

- Clones the "*Country*" field and maps the cloned field to the "Phone" qualified lookup field – the value will be the non-qualifier for the Phone field.
- Creates a new source field called "*Email Address Type*" which is mapped to the "E-Mail Address" qualified lookup field. The value "NULL" from the new source field is mapped to the value "Standard Address" as a fixed address type for all the imported addresses (since the source does not have an address-type field).
- Maps the "Number" field to the "[Remote Key]" destination field.
- Creates a new source field called "*Status*" which is mapped to the "Status" flat lookup field. The value "NULL" from the new source field is mapped to the value "Imported" as a fixed status for all newly imported records.
- Mapped the "State" source field to the "Region" and added all the Region values. When manually importing, the values need to be added.
- "NationalID" is mapped to "Search Term 1."
- "*GivenName*" is cloned and the clone is assigned to Name 1.
- "*GivenName*" and "Surname" are combined and mapped to "Customer [DF]."

Click on the "Import Status" tab and then on the import button ❗ in the toolbar to start the import process.

6.2.2 Loading Data from the COMPANYB Remote System

Repeat the same process as the previous step, only select "COMPANYB" as the Remote system and select the file "companyb_customers.xls" in the "Connect to Source" dialog box:

Then proceed to import the data the same way you did with COMPANYA. Be aware that you may have to add some of the field values, for example State.

7 Set Up Matching

Open Data Manager and login to the repository. You can see all the imported records. Notice they all have the value "Imported" in the Status field. There are also many similar records.

Switch to Matching Mode by clicking the Matching Mode icon in the toolbar: [icon].

7.1 Creating Transformations

Click on the "Transformations" tab. Right-click in the "Transformations" box and select "Add Transformation":
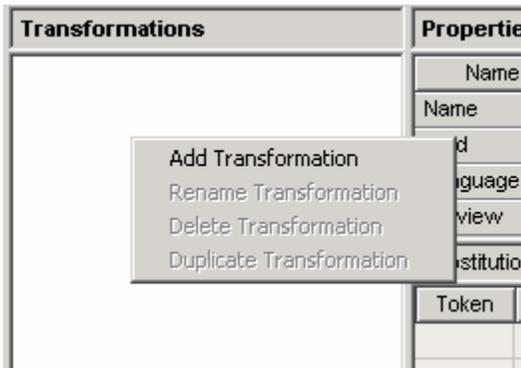
Name the new transformation "Remove nicknames and typos". Select "Name 1" as the Field and add the following substitutions:

Token	From	To
☑	Bobby	Robert
☑	Dick	Richard
☑	Bill	William
☑	Jon	John
☑	Chris	Christopher
☑	Ray	Raymond
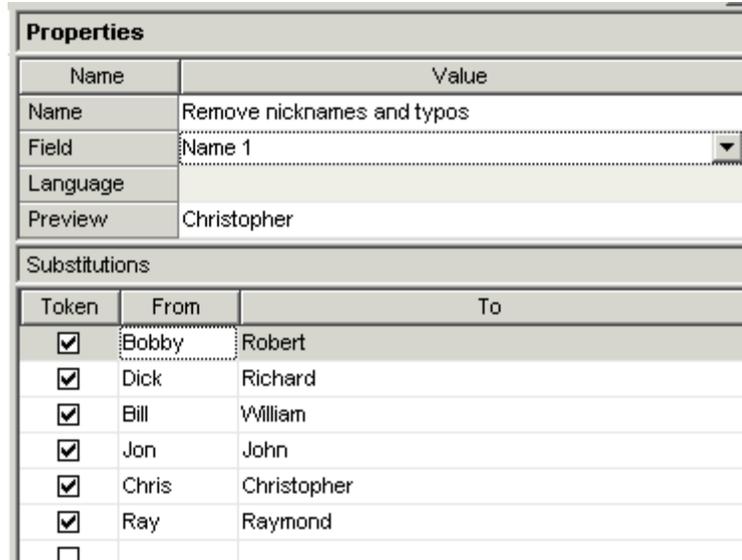
Seen on the right is the result.

Create another transformation named "Remove street suffixes" for the field "Street" with the following substitutions:

Token	From	To
☑	St	
☑	Street	
☑	Road	
☑	Rd	
☑	Lane	
☑	Drive	
☑	Dr	
☑	Avenue	
☑	Ave	

Seen on the right is the result.

7.2 Creating Rules

Click on the "Rules" tab. Right-click in the Rules section and select "Add Rule"
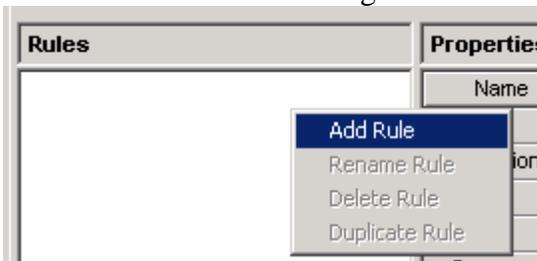
Name the new rule "Match on names", set Function to "Token Equals" and add the fields "Remove nicknames and typos [XFM]" and "Name 3":

The first field, "Remove nicknames and typos [XFM]" is the result of the transformation on the field "Name 1" which has the customer's first name. "Name 3" is the customer's last name. We ignore "Name 2" and "Name 4" because they store, according to the import map we used, the middle initial and maiden name which are of little interest for the de-duplication process.

Set the success value to "100" and the failure to "-40".

Properties	
Name	Values
Name	Match on names
Description	
Function	Token Equals
Fields	Remove nicknames and types [XFM]; Name 3
Success	100
Failure	-40
Undefined	0

Create a new rule with the following properties:

Properties	Value
Name	Match on National ID
Function	Equals
Fields	Search Term 1
Success	100
Failure	-50
Undefined	20

Properties	
Name	
Name	Match on National ID
Description	
Function	Equals
Fields	Search Term 1
Success	100
Failure	-50
Undefined	20

Create a third rule with the following properties:

Properties	Value
Name	Match on street
Function	Token Equals
Fields	Remove street suffixes [XFM]
Success	80
Failure	-20
Undefined	0

Properties	
Name	
Name	Match on street
Description	
Function	Token Equals
Fields	Remove street suffixes [XFM]
Success	80
Failure	-20
Undefined	0

7.3 Creating Strategies

Lastly, we will define the high and low threshold for which you think a match has occurred. Switch to the "Strategies" tab. Right-click in the Strategies box and select "Add Strategy"

Name the new strategy "Name match", include the "Match on names" rule, and set the high and low thresholds to 60 and 30, respectively.

Create another strategy named "Name and street match". Include the "Match on names" and "Match on street" rules, and set the high and low thresholds to 100 and 40, respectively:

Create a third strategy named "Match on ID". Include the rule "Match on ID" and set the high and low thresholds to 90 and 0, respectively:

SAP BEST PRACTICES AND SERVICE SOLUTION MANAGEMENT Page 23

7.4 Run the Matching Strategies and Find Duplicates

We will next utilize the Strategies that were created in the previous step.

Click on the "Matches tab". Select all records in the upper right pane, right-click, and select Match Records ▶ Selected vs. Selected ▶ Name match.

We will look at a specific record, the Customer "Dick Cantu" and see it has a High match to another record despite the difference in the first name – this is because of the transformation rules. You see the record has several matches, but only one high.

Records

	[Count]	[Level]	[Score]	Id	Customer
	3	Low	33	774	Derek Marquez
	0	None	-40	1320	Dewayne Embree
	2	Low	33	1032	Diane Hutcheson
	2	Low	33	1416	Diane Lemons
	2	Low	33	969	Diane Mellott
	1	Low	33	1428	Dianne Stark
▶	11	High	100	1083	Dick Cantu
	0	None	-40	1215	Dominic Altamirano
	0	None	-40	920	Dominique Proulx
	3	High	100	1076	Don Barnes

Matches | Merge | Transformations | Rules | Strategies | Workflows | Search Selection

	[Include]	[Level]	[Score]	[Match on r	Id	Customer	
▶	☐	High	100	100	825	Richard Cantu	Imported
	☐	Low	33	33	622	Richard Thomas	Imported
	☐	Low	33	33	895	Richard Lamoureux	Imported
	☐	Low	33	33	988	Richard Watterson	Imported
	☐	Low	33	33	1015	Richard Yazzie	Imported
	☐	Low	33	33	1091	Richard Watterson	Imported
	☐	Low	33	33	1235	Richard Tompkins	Imported
	☐	Low	33	33	1319	Richard Morris	Imported
	☐	Low	33	33	1408	Richard White	Imported
	☐	Low	33	33	1424	Richard Colon	Imported
	☐	Low	33	33	1469	Richard Hampton	Imported

To narrow the results, right-click the record, Dick Cantu, and select "Match Records → Selected vs. Results → Name and street match" to run a more elaborate matching strategy on the results of the previous match. We now see that this filters out all of the low-class matches and leaves you with only one record.

Records

	[Count]	[Class]	[Score]	Id	Customer
				918	Diane Lemons
				471	Diane Mellott
				930	Dianne Stark
▶	1	High	180	584	Dick Cantu
				716	Dominic Altamirano
				422	Dominique Proulx
				577	Don Barnes

Matches	Merge	Transformations	Rules	Strategies	Workflows	Search Sele

	[Include]	[Class]	[Score]	[Match on n	[Match on s	Id	
▶	☐	High	180	100	80	327	Richard Cantu

Click on the "Include" checkbox for the record with Level High and then on the "Merge" tab. Right-click on the header of the second record and select "Set All":

Matches	**Merge**	Transformations	Rules	Strategies	Workflows	Search Selections	

|	**[Merged Record]**	▶ Dick Cantu ◀	Richard Cantu	
Customer		Dick Cantu	Richard Cantu	Hide
Status		Imported	Imported	Unhide ▶
Account group				Sort ▶
Title				Width ▶
Name 1		Dick	Richard	Set All (Equal)
Name 2		B	B	Set All (Equal or NULL)
Name 3		Cantu	Cantu	**Set All**
Name 4		Andrews	Andrews	Set with Non-NULLs Only
Search Term 1		243-85-9002	243-85-9002	Set NULLs Only
Search Term 2				
Street		3525 Diamond Street	3525 Diamond Stree	Clear All
House Number				
City Postal Code		28801	28801	
City		Asheville	Asheville	
Country		United States	United States	
Region		NC	NC	
Address Time Zone				
Building				

This will copy all the information to the new record [Merged Record].

Right-click anywhere in the lower part of the screen and select "Merge Records":
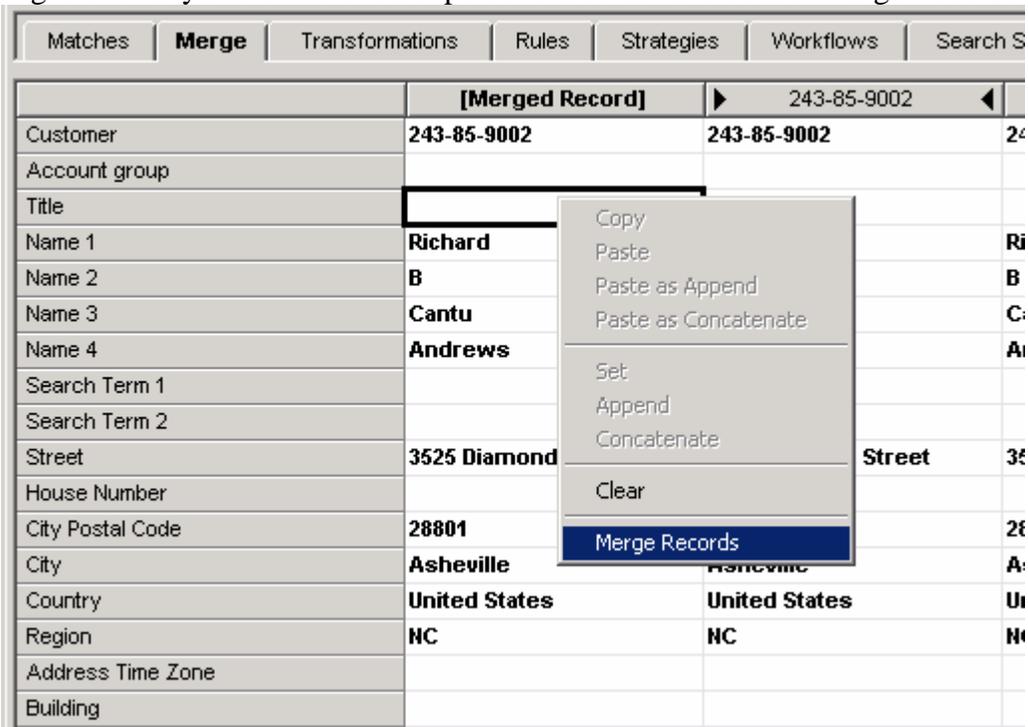
Answer "Yes" in the popup window. MDM will now merge the two records. Switch back to Record Mode by clicking the record mode button in the toolbar and find the merged record under Richard Cantu. Right-click the record and select "Edit Key Mappings" to see that the record is in fact the result of the merger of the two original records which came from two different sources:
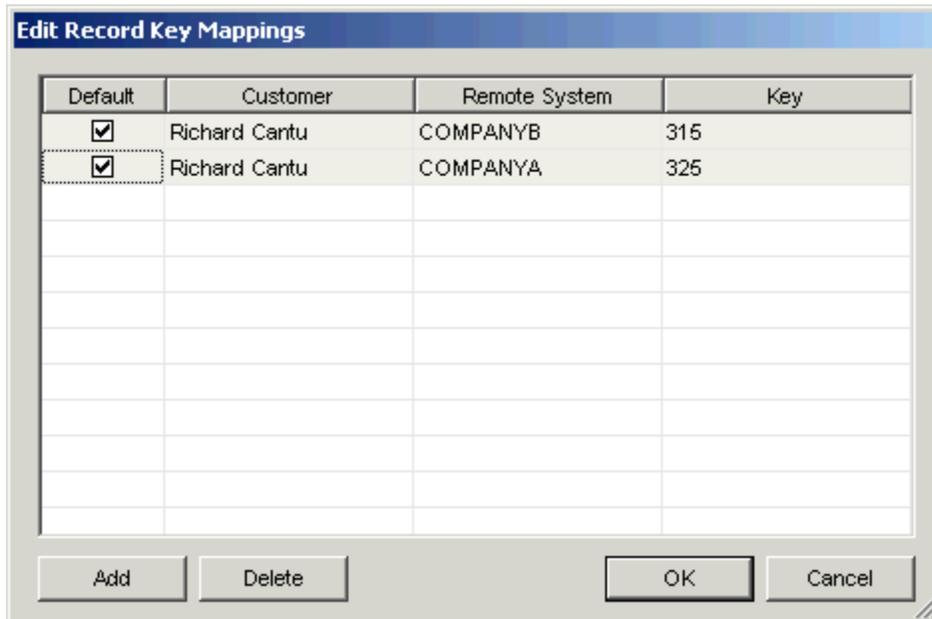
You can now proceed to identify and merge other matching records. Many of the keys are the same, but others are not.